

# Topics in Natural Language Processing Section 01

## CS 273

Fall 2023 3 Unit(s) 08/21/2023 to 12/06/2023 Modified 08/19/2023

### Contact Information

Instructor(s): Faranak Abri

Office Location: MQ 213

Telephone: 4089244844

Email: faranak.abri@sjsu.edu

Office Hours: Mondays-Wednesdays 4:30PM-5:30PM (Zoom or in-person by appointment)

Class Days/Time: Mondays – Wednesdays 3:00-4:15PM

Classroom: Science Building 311

### \* Classroom Protocols

- *Students are expected to assist in maintaining a classroom environment that is conducive to learning. Inappropriate behavior in the classroom that leads to the distraction of others shall not be tolerated under any circumstances.*
- *Instruction will begin at or within several minutes of the official published start time for the course. Please make sure that cell phones, beepers, and texting devices are turned off during the entire scheduled class time. Excessive audible discussions with fellow students are prohibited so that others are not disturbed. If any subject matter is not understood, please do not hesitate to ask for clarification. If an extended response is necessary to remove doubts, then a request to follow up outside of scheduled classroom instruction time might be made.*
- *Per [University Policy S12-7\(Links to an external site.\)](#), course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without permission. Students may not publicly share or upload instructor-generated material for this course such as exam questions, lecture notes, or homework solutions without the instructor's consent. This includes unauthorized recording or posting of recordings of lectures. Students who record, distribute, or post these materials will be referred to the Student Conduct and Ethical Development office. These policies are designed to protect student privacy and ensure academic integrity.*
- *If a student is caught cheating on a homework assignment, the student will receive a 0 on that assignment. If a student is caught cheating on an exam, the student will receive an F.. The instructor must report any incidents of cheating or plagiarism to the University per [University Policy F15-7\(Links to an external site.\)](#).*

### Program Information

Diversity Statement - At SJSU, it is important to create a safe learning environment where we can explore, learn, and grow together. We strive to build a diverse, equitable, inclusive culture that values, encourages, and supports students from all backgrounds and experiences.

### Course Learning Outcomes (CLOs)

Upon successful completion of this course, students will be able to:

- *CLO1 Design regression and classification learners (models) and evaluate the goodness of their designed Models based on overfitting/underfitting and bias/variance.*
- *CLO2 Process textual data, and design/evaluate ML models to classify textual data*
- *CLO3 Analyze the main applications of NLP and implement different NLP projects.*
- *CLO4 Write technical reports about their projects and present their findings.*

## Course Materials

---

### Textbooks (Optional)

*Pattern recognition and machine learning. Christopher M. Bishop; 2006*

*Machine Learning. Sergios Theodoridis; 2015*

*Speech and Language Processing (An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition). Dan Jurafsky and James H. Martin.*

*Foundations of Statistical Natural Language Processing. Christopher Manning and Hinrich Schütze*

*Other online resources*

## Course Requirements and Assignments

---

- *Two exams 40% = First exam 20% + Second exam 20%,*
- *Three projects 60% (Each project includes one or more subprojects, written reports in LaTeX format) = First project 10% + Second Project 20% + Capstone project (required presentation) 30%*
  - *grading rubrics for Project 1 and Project 2: Implementing all mentioned steps (0.6) and clear report (0.4)*
  - *Capstone project: Students must choose a project for their final project that will require an adequate amount of teamwork and then carry it out. The capstone project is a mid-scale project that can be composed of smaller projects, and students work in groups to accomplish their project.*
  - *The Capstone project proposal is sent to the instructor for approval before starting to work on it. The proposal includes the description of the NLP project, putting it out in one page's worth of detail.*

*Briefly describe the dataset(s) you plan to work on.*

*List the steps you want to take for implementation.*

*Establish what the project's results will be.*

- *grading rubrics for Capstone project: Project proposal (0.1), implemented project (0.5), Report (0.2), presentation (0.2)*
- *Bonus points = Presenting research work for specific topics and presenting first and second projects after the instructor's approval.*

*"Per [University Policy S16-9](#), success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course-related activities, including but not limited to internships, labs, and clinical practices. Other course structures will have equivalent workload expectations as described in the syllabus."*

### Final Examination or Evaluation

*This is a project-based course, and the final examination is done by the capstone project and the written report in LaTeX format and the presentation (30%).*

## ✓ Grading Information

---

*Extra credit questions may be included in assignments and exams. For late submissions, grade deductions will be considered.*

*A plus = 97% or higher*

*A = 93% up to 97%*

*A minus = 90% to 93%*

*B plus = 87% to 90%*

*B = 83% to 87%*

*B minus = 80% to 83%*

*C plus = 77% to 80%*

*C = 73% to 77%*

*C minus = 70% to 73%*

*D plus = 67% to 70%*

*D = 63% to 67%*

*D minus = 60% to 63%*

*F = 0% to 60%*

## University Policies

---

Per [University Policy S16-9 \(PDF\)](http://www.sjsu.edu/senate/docs/S16-9.pdf) (<http://www.sjsu.edu/senate/docs/S16-9.pdf>), relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on the [Syllabus Information](https://www.sjsu.edu/curriculum/courses/syllabus-info.php) (<https://www.sjsu.edu/curriculum/courses/syllabus-info.php>) web page. Make sure to visit this page to review and be aware of these university policies and resources.

## Course Schedule

---

# CS 273 / Topics in Natural Language Processing, Fall 2023, Course Schedule

Week/session	Date	Topics, Readings, Assignments, Deadlines
Week 1	Aug 21 , Aug 23	Introduction
Week 2	Aug 28 , Aug 30	Review of evaluation of ML/NLP models, working with overleaf and LaTeX
Week 3	<del>Sep 4</del> Labor Day - Campus Closed (L), Sep 6	Review of evaluation of ML/NLP models

Week 4	Sep 11 , Sep 13	Text Preprocessing: regular expressions, stemming, lemmatization, stop words, text Normalization, tokenization
Week 5	Sep 18 , Sep 20	Text Preprocessing: regular expressions, stemming, lemmatization, stop words, text Normalization, tokenization
Week 6	Sep 26 , Sep 27	Edit Distance, word and sentence similarity, and cosine similarity
Week 7	Oct 2 , Oct 4	Word Embedding, Word vectors, Word senses <a href="#">First Exam (Oct 4)</a>
Week 8	Oct 9 , Oct 11	N-grams language models, evaluation, and smoothing.
Week 9	Oct 16 , Oct 18	Spelling correction
Week 10	Oct 23 , Oct 25	Text classification and Naïve Bayes, RNN, LSTM
Week 11	Oct 30 , Nov 1	Sentiment classification, sentiment lexicons, Named Entity Recognition
Week 12	Nov 6 , Nov 8	Corps-based chatbots rule-based chatbots, properties of human conversations, keyword search (Information Extraction)
Week 13	Nov 13 , Nov 15	Pre-Trained models, large language models, and fine-tuning
Week 14	Nov 20 , <del>Nov 22</del> Non-Instructional Day	<a href="#">Second Exam (Nov 20)</a>
Week 15	Nov 27 , Nov 29	Capstone project presentations
Week 16	Dec 4 , Dec 6	Capstone project presentations

No Final Exam

Other important dates:

Sep 15: Last Day to Drop Classes without a "W" Grade

-----

Fall 2023 calendar:

[https://www.sjsu.edu/provost/docs/Academic\\_Calendar-AY2023-24.pdf](https://www.sjsu.edu/provost/docs/Academic_Calendar-AY2023-24.pdf)